NOTES AND UNIQUE PHENOMENA

PREDICTING YIELD LOSS IN INDETERMINATE SOYBEAN FROM POD DENSITY USING SIMULATED DAMAGE STUDIES

J. W. Singer,* R. W. Malone, D. W. Meek, and D. Drake

Abstract

Developing relationships between seed yield and pod density can be useful for predicting yield loss in soybean [*Glycine max* (L.) Merr.] damaged by deer (*Odocoileus virginianus*). The objectives of this research were to (i) develop a modeling tool using differences between biomass removal treatments and controls for pod density and seed yield to quantify yield loss and (ii) assess the tool using double cross-validation. Model development using linear and polynomial exponential (PE) equations was accomplished using 1998–2001 data from studies examining different biomass removal treatments, varieties, and row spacings. The PE model had a slightly higher coefficient of determination (R^2 = 0.93) than the linear model (R^2 = 0.92). Double cross-validation of both models produced strong relationships with high coefficients of determination and predictive ability; however, the model performance statistics indicated that the PE model had higher coefficients of determination, lower mean bias error, and more robust slope estimates than the linear model. Depending on the end-user, the simplicity of the linear model should be carefully considered in weighing the benefits of each tool. Nevertheless, these approaches provide robust tools that are not sensitive to moderate abiotic fluctuations, varying cultural practices, and a wide range of temporal biomass removal. Validating the relationship using additional data should be the next step before implementation.

Pod number and yield in defoliated determinate soybean may be correlated (Board and Harville, 1993; Board and Tan, 1995). Goli and Weaver (1986) evaluated complete defoliation of late-planted determinate and indeterminate soybean cultivars at R4 (Fehr and Caviness, 1977), R5, and R6 and concluded that yield loss was primarily attributed to a reduction in the number of pods per plant and that the indeterminate growth habit did not provide an advantage in buffering seed yield. However, the pod response to biomass removal in indeterminate soybean under varying biomass removal treatments and cultural practices has not been thoroughly evaluated.

J.W. Singer, R.W. Malone, and D.W. Meek, USDA-ARS Natl. Soil Tilth Lab., 2150 Pammel Drive, Ames, IA 50011; and D. Drake, Dep. of Ecol., Evolution, and Nat. Resour., Cook College, Rutgers Univ., 80 Nichol Ave., New Brunswick, NJ 08901. Received 28 March 2003. *Corresponding author (singer@nstl.gov).

Published in Agron. J. 96:584–589 (2004).

Timing, intensity, and frequency of biomass removal affects soybean yield. Yield reduction is less sensitive to biomass removal during vegetative growth because soybean can develop new leaf area that can compensate for temporarily reduced assimilatory capacity. Singer (2001) reported that yield reductions in indeterminate soybean from removing the top third of the plant at V5 were less than biomass removed at R4. Although indeterminate soybean vegetative growth occurs until R5, creating new leaf area that increases assimilatory capacity directly competes with reproductive sink demand. Fehr et al. (1977) reported that yield of determinate cultivars was affected more than indeterminate from 100% defoliation when defoliation occurred from R2 through R6. However, average yield loss from half-plant cutoff was similar for determinate (33%) and indeterminate (34%) cultivars, but there was a significant interaction with growth stage (Fehr et al., 1977).

White-tailed deer are among the most identifiable types of wildlife in North America and provide many aesthetic, recreational, economic, and ecological benefits. However, deer can also cause negative economic and ecological impacts in areas where they are overabundant locally or regionally. The agricultural community experiences many of the negative economic impacts. For example, nationwide, deer have been recognized to cause more damage to agricultural crops than any other vertebrate wildlife species (Conover and Decker, 1991), costing farmers more than an estimated $100 million each year (Conover, 1997, 1998). The greatest agricultural damage by deer generally occurs in the northeastern and north-central United States where at least 41% of producers reported damage (Wywialowski, 1994).

Most estimates regarding economic damage to agricultural crops from deer are based on perceptions of agricultural producers (Conover, 1994) and wildlife professionals (Conover and Decker, 1991). Documented validation of crop depredation to support or refute perceptions is needed so that damage management policies and strategies can be enacted, if necessary, to reduce/eliminate conflicts between agriculture and deer. As of 1994, 18 states in the USA had established compensation programs for damage caused by ungulates (Wagner et al., 1997).

Many of the strategies currently available for quantifying wildlife depredation to agriculture can be labor intensive and costly, especially when assessing damage to row crops (Wisconsin Wildlife Damage Abatement and Claims Program, 2000). A common method for quantitative assessment of deer depredation to crops is to construct exclosures randomly placed within a field. Vecellio et al. (1994) constructed 5-m-long by approximately 5-m-wide by 1.8-m-high exclosures for measuring

Abbreviations: CI, confidence interval; OLS, ordinary least squares; PE, polynomial exponential; RMSE, root mean squared error.

yield loss in corn (*Zea mays* L.) and considerably smaller areas for measuring wheat (*Triticum aestivum* L.) losses. Comparisons were made between yield inside and outside the exclosure to determine yield loss. Although establishing exclosures requires time and materials, processing samples for yield determination is the least efficient component of this methodology and requires a substantial amount of time for each exclosure. Consequently, a relatively quick and accurate methodology to quantify white-tailed deer damage to soybean would be of great benefit. The objectives of this research were to (i) develop a modeling tool using differences between biomass removal treatments and controls for pod density and seed yield to quantify yield loss and (iii) assess the tool using double cross-validation.

Materials and Methods

The data presented were collected from 1998 through 2001 in two separate 2-yr field studies. Both studies were conducted on a Quakertown silt loam (fine-loamy, mixed, mesic Typic Hapludult) at the Rutgers University Snyder Research and Extension Farm near Pittstown, NJ (40°30′ N, 75°00′ W). The first study was conducted from 1998–1999 and used indeterminate soybean 'Golden Harvest H-1357RR' planted using conventional tillage on 4 June and 27 May 1998 and 1999, respectively, at 494 000 seeds ha^{-1} in narrow (18 and 20 cm in 1998 and 1999, respectively) and wide (76 cm) rows. In the second study, indeterminate soybean 'Pioneer Brand 93B53' and 'Agway Brand APK394NRR' were seeded using no-till on 16 and 21 May in 2000 and 2001, respectively, at 518 700 seeds ha^{-1} in narrow (20 cm), intermediate (41 cm), and wide (76 cm) rows. Fertilizer was applied according to soil test recommendations in both experiments. In the first study, soybean was evaluated each year following a rye (*Secale cereale* L.) cover crop. In the second study, soybean was evaluated each year after corn. Pre-emergence herbicides were used for weed control in both studies.

Experimental design in the first study was a randomized complete block in a split-plot arrangement with three replications in 1998 and four in 1999. Main plot was narrow vs. wide row spacing. Subplot treatments were a control and biomass removal at V5, R1, R4, and all combinations for a total of seven biomass removal treatments. Subplot size was 15 m^2 for narrow rows and 23 m^2 for wide rows. Biomass removal was accomplished using hedge clippers to remove top growth by measuring the height of each treatment and removing approximately 30% of the average height. Growth stages were determined using the control plot as a reference.

In the second study, the experimental design was a randomized complete block in a split-split plot arrangement with four replications. Main plot was variety, subplot was row spacing, and sub-subplot was biomass removal. Soybean biomass was removed at V1 + V3 + V6, V6 + R1, R1 + R4 + R6, V1 + V3 + V6 + R1 + R4 + R6, and a control. Sub-subplot size was 4 m^2 for the narrow and intermediate row spacing and 6 m^2 for the wide row spacing. Biomass removal was accomplished using scissors to remove top growth by measuring the height of each treatment and removing approximately 30% of the average height. Growth stages were determined using the control plot as a reference.

Ten and 15 plants per plot, in 1998–1999 and 2000–2001, respectively, were randomly sampled after physiological maturity to determine pod number and seed yield. Pod number and seed yield were converted to an area basis using harvest stand counts. Soybean seed were dried in a forced-air oven at 70°C for at least 72 h and weighed to determine yield.

Differences between the control (no biomass removal) and biomass removal treatments were calculated for each replication. The independent variable was calculated as the difference between a treatment and the control for pod density. The dependent variable was calculated as the difference between a treatment and the control for seed yield. Regression analysis was conducted using treatment means for the difference between seed yield and pod density. Mean seed yield for all observations (n = 352) from 1998–2001 was 433 g m^{-2} (median = 406 g m^{-2}) with a standard error of 12. Mean pod density was 1353 pods m^{-2} (median = 1282 pods m^{-2}) with a standard error of 32.

Ordinary least squares (OLS) methods were employed instead of measurement error methods because the goal of the modeling was prediction and the measurement error in pod density was small (reliability ratio, κ = 0.945; Fuller, 1987). The OLS procedures were weighted to correct for mild heteroscedasticity. Weighting was accomplished using the inverse variance weight model with the variance proportional to the prediction ($\sigma^2 = \hat{y}$ or $\sigma^2 = \hat{y}^{1/2}$, where $\hat{y}$ = predicted value). Treatment means were split at random into two sets with equal numbers of observations (Montgomery and Peck, 1982). Two models were developed for the full data set and each split data set. To select the best model, double cross-validation assessment was performed to assess the predictive performance for each model from each split set (Kaspar et al., 2003). Model performance evaluation criteria were mean bias error (Fox, 1981), slope, root mean squared error (RMSE), and the coefficient of determination (R^2). Multiple slope estimates were evaluated using models with and without the intercept through the origin. All analyses were performed using the Statistical Analysis System (SAS Inst., 2001).

Results and Discussion

Results from a simple linear regression of pod density difference on yield difference were highly significant using all the data from 1998 through 2001 (n = 76; R^2 = 0.92). The PE model also provided a strong relationship between pod density difference and seed yield (n = 76; R^2 = 0.93). The linear model equation using OLS regression is

$$y = -50.1 + 0.421x \quad [1]$$

where y = yield difference and x = pod density difference. The 95% confidence interval (CI) for the linear model is presented in Fig. 1A. The model using the PE is

$$y = -705 + \exp(6.5 + 0.000558x - 0.000000572x^2) \quad [2]$$

The 95% CI for the PE model is presented in Fig. 1B. Both models perform well, but the PE has slightly better interpolation than the linear model. The 95% CI band increases as absolute differences increase because of the larger error associated with larger differences.

These relationships indicate that pod density can be used to explain a high degree of the variability in yield across variable precipitation levels with single biomass removal events or repeated biomass removal during vegetative and reproductive growth stages. Although precipitation at the experimental site was below average in 1998 and 1999 and close to average in 2000 and 2001, seed yields were similar in 1998 and 2001 and for 1999

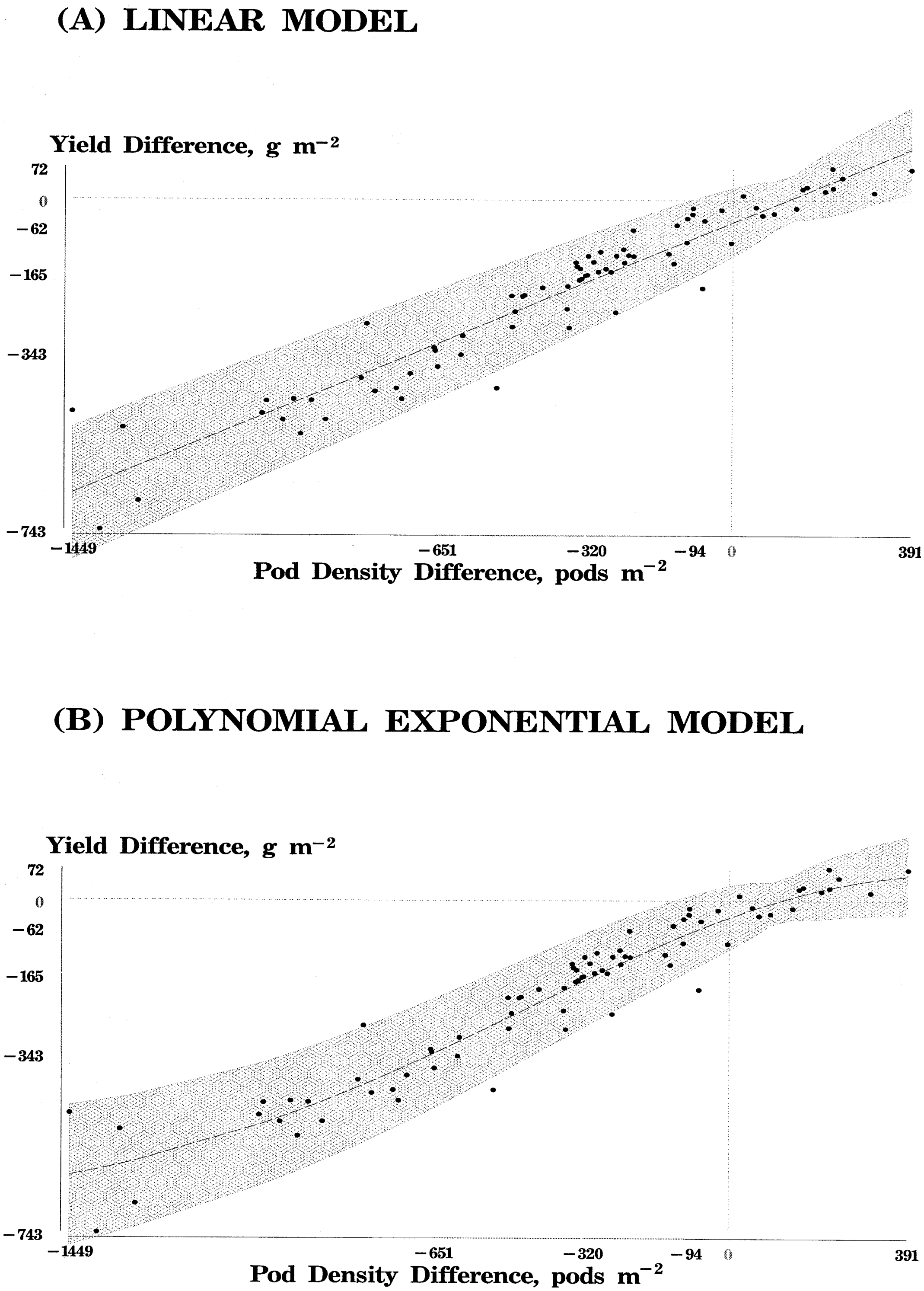

Fig. 1. Relationship between yield difference and pod density difference. Yield difference is between biomass removal treatments and the control (no biomass removal). The shaded area represents a 95% confidence interval. All axes are divided into quartiles.

and 2000 (data not presented). Caviness and Thomas (1980) reported that yield reduction of defoliated soybean was similar under irrigated and nonirrigated conditions, even in extremely and moderately dry growing seasons, and that reductions in pod number appeared to be the yield component primarily responsible for the yield losses. Consequently, the pod density–seed yield difference relationship appears quite robust across growing seasons with different amounts and timing of precipitation.

Our data also provide a strong relationship across row spacings. We evaluated biomass removal in row spacings ranging from 18 to 76 cm. Midwest research documented a linear decrease in seed yield and pod

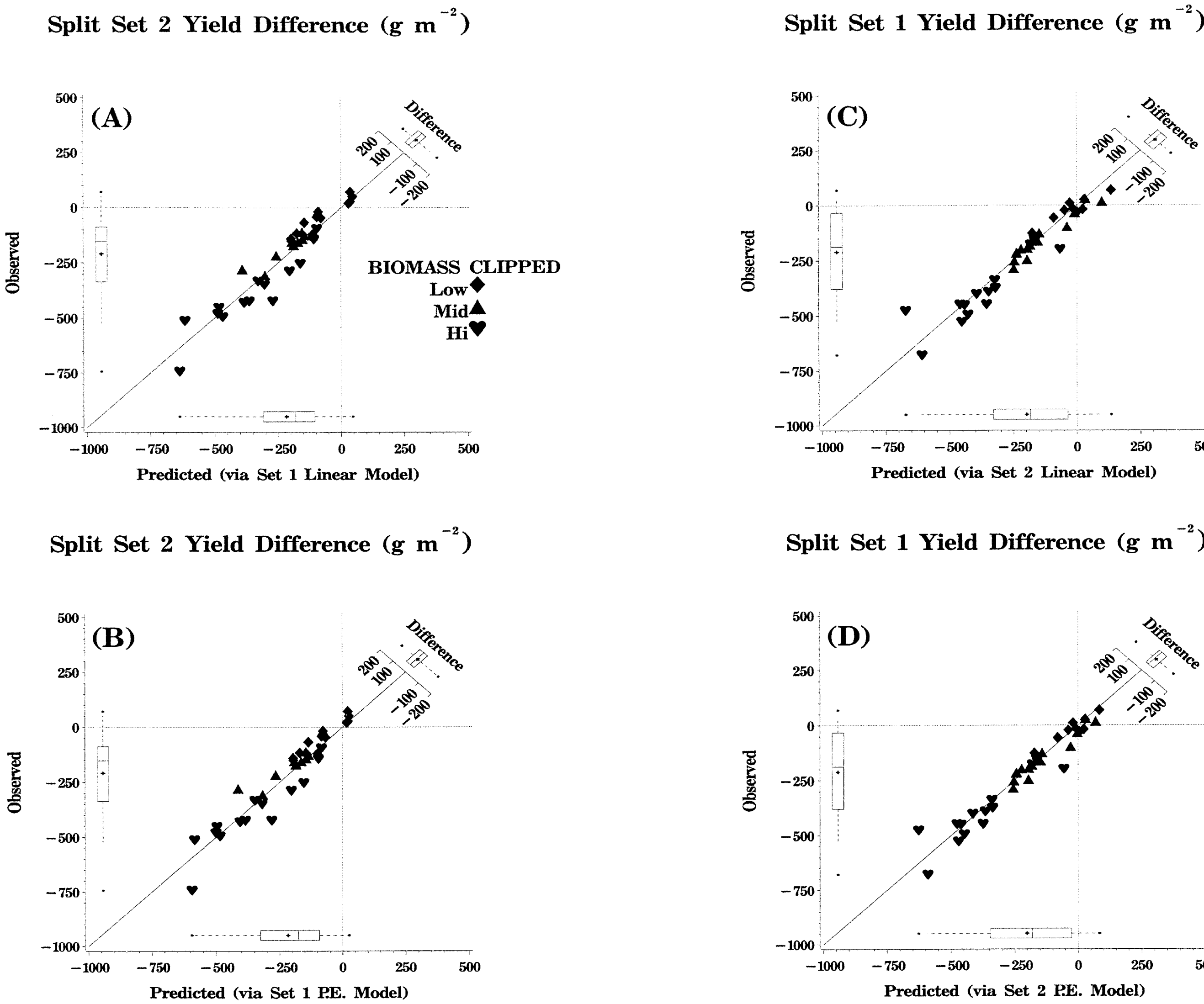

Fig. 2. Berg (1992) plot of observed vs. predicted yield difference for double cross-validation of linear and polynomial exponential models using randomly split data sets. Yield difference is between mean biomass removal treatments and the control (no biomass removal). Low biomass removal treatments include V5, R1, V5 + R1, and V6 + R1; medium treatments include R4, V5 + R4, R1 + R4, and V1 + V3 + V6; and high treatments include V5 + R1 + R4, R1 + R4 + R6, and V1 + V3 + V6 + R1 + R4 + R6.

number per plant as row width increased in indeterminate soybean (Bullock et al., 1998) while Board et al. (1992) under Louisiana conditions reported that the main factors responsible for increased yields in narrow rows were greater fertile node production and increased pod number per fertile node. Clearly, the relationship we are reporting between pod density difference and seed yield difference is not as sensitive to row spacing.

Additionally, Fehr et al. (1977) reported that cultivars within location responded similarly to defoliation and half-plant cutoff for most of the measured characters and that the percentage change was similar. They concluded, however, that the determinateness of a cultivar should be considered when assessing yield reduction from defoliation and half-plant cutoff during reproductive development. Consequently, inference from the relationship between pod density and seed yield that we are reporting should be limited to indeterminate cultivars, particularly when biomass removal occurs during reproductive development.

The second objective of this study was to assess the predictive performance of the tools. The entire data set was randomly split into two equal separate sets. By randomly splitting the data, we averaged across years with different amounts and timing of precipitation and cultural practices. The 2000–2001 experiment had two varieties (MG 3.5 and 3.9, respectively), three row spacings, and more frequent biomass removal than the 1998–1999 study, which used conventional tillage, a different MG 3.5 variety, similar narrow and wide row spacings, and similar intensity, but less frequent, biomass removal.

Although abiotic conditions and cultural practices were different, the random split of the data into two sets provided a high degree of prediction. The Linear Model Set 2, which was predicted using the Set 1 data (Fig. 2A), had a coefficient of determination slightly lower than the full data set ($n = 38$, $R^2 = 0.91$, RMSE = 56). Not surprisingly, the most extreme yield loss occurred in the treatment where biomass was removed at three vegetative and reproductive growth stages. The PE Model Set 2, which was predicted from the Set 1 data (Fig. 2B), had a coefficient of determination slightly lower than the full data set ($n = 38$, $R^2 = 0.91$, RMSE = 56). Predictive ability for both models was high although mean bias error was higher for the linear (8.03 ± 9.15 g m^{-2}) compared with the PE model (7.91 ± 9.17 g m^{-2}). Both mean bias error estimates were systematically overestimating yield loss difference, but these bias estimates were not significant ($t = 0.88$ and 0.86 for the linear and PE models, respectively). Using a weighted OLS model, the t statistic for the linear (1.4) and PE (1.7) models indicated that the intercepts were not significant, and both slope estimates included 1 in the 95% CI.

The cross-validation using Set 2 data to predict Set 1 values for the linear model (Fig. 2C) had a coefficient of determination slightly lower than the full data set ($n = 38$, $R^2 = 0.91$, RMSE = 55). One data point in the high category was not predicted well. This data point was from a treatment that included biomass removal at R1 + R4 + R6 growth stages in 2001, a year with favorable precipitation. The PE model using Set 2 to predict Set 1 (Fig. 2D) had a slightly higher coefficient of determination ($n = 38$, $R^2 = 0.92$, RMSE = 48) than the first assessment and slightly lower than the full model using the entire data set. The PE model had lower mean bias error (-9.29 ± 7.68 g m^{-2}) than the linear model (-12.08 ± 8.78 g m^{-2}). Using the Set 2 data to predict Set 1 values systematically underestimated yield loss for both models although neither bias term was significant. Both slope estimates included 1 in the 95% CI using a weighted OLS model with the regression line through the origin. The PE model performed better than the linear model for the Set 1 cross-validation using Set 2 for prediction. Depending on the end-user's needs for these tools, the simplicity of the linear model should be carefully considered in weighing the benefits of each tool.

Typically, wildlife biologists and crop scientists install multiple exclosures in a field and compare yield inside the exclosure to yield outside the exclosure to determine yield loss due to wildlife, in this instance, white-tailed deer. Obtaining yield outside the exclosure usually, but not always, involves a mechanical harvest system. Determining yield inside the exclosure, however, is extremely labor intensive. The crop is hand-harvested and transported back to the research farm, and with soybean, a small-plot thresher is used to obtain seed yield. Approximate time for this process is about 6 to 8 h per field (Dr. David Drake, personal communication, 2003). Our model eliminates the need to harvest the crop and merely relies on counting the number of pods per unit area inside the exclosure and in representative areas outside the exclosure. Even though exclosures are still required for control data, this process can reduce the time it takes to quantify yield loss by about 85%, or to about 1 h per field. This is a major time savings and will increase the efficiency of quantifying soybean yield loss.

Conclusions

A strong relationship exists in soybean between pod density difference and seed yield difference using simulated biomass removal techniques. Cross-validation of both models also produced strong relationships with high coefficients of determination and predictive ability within 5 to 6% of actual values. Consequently, both the linear and PE models can be used as tools to predict yield loss in soybean damaged by deer, but the PE model performs somewhat better. This approach to quantify yield loss does not totally eliminate the need for representative exclosures to be established for reference soybean yield in areas where high deer densities occur. Nevertheless, implementing this approach can increase the efficiency of quantifying indeterminate soybean yield loss to deer damage on a larger scale and provide an objective methodology. Validating the relationship using additional data should be the next step before implementation.

References

Berg, R.L. 1992. First place. Best presentation of data-monochrome. p. 1521–1527. *In* Proc. Annu. SAS Users Group Int. Conf., 17th, Honolulu, HI. 12–15 Apr. 1992. SAS Inst., Cary, NC.

Board, J.E., and B.G. Harville. 1993. Soybean yield component responses to a light interception gradient during the reproductive period. Crop Sci. 33:772–777.

Board, J.E., M. Kamal, and B.G. Harville. 1992. Temporal importance of greater light interception to increased yield in narrow-row soybean. Agron. J. 84:575–579.

Board, J.E., and Q. Tan. 1995. Assimilatory capacity effects on soybean yield components and pod number. Crop Sci. 35:846–851.

Bullock, D., S. Khan, and A. Rayburn. 1998. Soybean yield response to narrow rows is largely due to enhanced early growth. Crop Sci. 38:1011–1016.

Caviness, C.E., and J.D. Thomas. 1980. Yield reduction from defoliation of irrigated and non-irrigated soybeans. Agron. J. 72:977–980.

Conover, M.R. 1994. Perceptions of grass-roots leaders of the agricultural community about wildlife damage on their farms and ranches. Wildl. Soc. Bull. 22:94–100.

Conover, M.R. 1997. Monetary and intangible valuation of deer in the United States. Wildl. Soc. Bull. 25:298–305.

Conover, M.R. 1998. Perceptions of American agricultural producers about wildlife on their farms and ranches. Wildl. Soc. Bull. 26: 597–604.

Conover, M.R., and D.J. Decker. 1991. Wildlife damage to crops: Perceptions of agricultural and wildlife professionals in 1957 and 1987. Wildl. Soc. Bull. 19:46–52.

Fehr, W.R., and C.E. Caviness. 1977. Stages of soybean development. Spec. Rep. 80. Iowa Agric. Exp. Stn., Ames.

Fehr, W.R., C.E. Caviness, and J.J. Vorst. 1977. Response of indeterminate and determinate soybean cultivars to defoliation and half-plant cut-off. Crop Sci. 17:913–917.

Fox, D.G. 1981. Judging air quality model performance: A summary of the AMS workshop on dispersion model performance. Bull. Am. Meteorol. Soc. 62:599–609.

Fuller, W.A. 1987. Measurement error models. John Wiley & Sons, New York.

Goli, A., and D.B. Weaver. 1986. Defoliation responses of determinate and indeterminate late-planted soybeans. Crop Sci. 26:156–159.

Kaspar, T.C., T.S. Colvin, D.B. Jaynes, D.L. Karlen, D.E. James, D.W. Meek, D. Pulido, and H. Butler. 2003. Relationship between six years of corn yields and terrain attributes. Precis. Agric. 4: 87–101.

Montgomery, D.C., and E.A. Peck. 1982. Introduction to linear regression analysis. John Wiley & Sons, New York.

SAS Institute. 2001. SAS STAT Software: Changes and enhancements, Release 8.2. Publ. 58309. SAS Inst., Cary, NC.
Singer, J.W. 2001. Soybean light interception and yield response to row spacing and biomass removal. Crop Sci. 41:424–429.
Vecellio, G.M., R.H. Yahner, and G.L. Storm. 1994. Crop damage by deer at Gettysburg Park. Wildl. Soc. Bull. 22:89–93.
Wagner, K.K., R.H. Schmidt, and M.R. Conover. 1997. Compensation programs for wildlife damage in North America. Wildl. Soc. Bull. 25:312–319.
Wisconsin Wildlife Damage Abatement and Claims Program. 2000. Technical manual field handbook. Wisconsin Dep. of Nat. Resour., Madison.
Wywialowski, A. 1994. Agricultural producers' perceptions of wildlife-caused losses. Wildl. Soc. Bull. 22:370–382.

New Books Received

Creating a Sustainable Future: Living in Harmony with the Earth. *Peter B. Kaufman, Christopher W. Coon, J.N. Govil, James E. Hoyt, Casey R. Lu, Barbara J. Madsen, and Sara L. Warber*. Studium Press, P.O. Box 722200, Houston, TX 77072. 2002. ISBN 1-930813-01-5.

Perspectives in World Food and Agriculture 2004. *Edited by Colin G. Scanes and John A. Miranowski*. Iowa State Press, 2121 State Ave., Ames, IA 50014-8300. 2004. ISBN 0-8138-2021-9. $69.99 (hard cover).

Pest Management in Citrus. *K.P. Srivastava and Y.S. Ahlawat*. Studium Press, P.O. Box 722200, Houston, TX 77072. 1999. ISBN 0-9656038-4-9.

Pest Management in Vegetables. Part 1. *K.P. Srivastava and Dhamo K. Butani*. Studium Press, P.O. Box 722200, Houston, TX 77072. 1998. ISBN 0-9656038-2-2.

Pest Management in Vegetables. Part 2. *K.P. Srivastava and Dhamo K. Butani*. Studium Press, P.O. Box 722200, Houston, TX 77072. 1998. ISBN 0-9656038-2-2.

Soil Fertility Decline in the Tropics with Case Studies on Plantations. *Alfred E. Hartemink*. CABI Publishing, 44 Brattle St., 4th Floor, Cambridge, MA 02138. 2003. ISBN 0-85199-670-1. $120.00 (hard cover).

Water Dynamics in Plant Production. *Wilfried Ehlers and Michael Goss*. CABI Publishing, 875 Massachusetts Ave., 7th Floor, Cambridge, MA 02139. 2003. ISBN 0-85199-694-9. $120.00 (hard cover).